Yee & Associates, P.C.

4100 Alpha Road
Suite 1100
Dallas, Texas 75244

Main No. (972) 385-8777
Facsimile (972) 385-7766

**RECEIVED
CENTRAL FAX CENTER**

APR 15 2005

Facsimile Cover Sheet

To: Commissioner for Patents for Examiner Laurie Anne Ries Group Art Unit 2176	Facsimile No.: 703/872-9306
From: Amelia Turner Legal Assistant to Stephen R. Tkacs	No. of Pages Including Cover Sheet: 21

Message:

Transmitted herewith:

- Transmittal Document; and
- Appeal Brief.

Re: Application No.: **09/838,428**
    Attorney Docket No: **AUS920010015US1**

Date: Friday, April 15, 2005

| **Please contact us at (972) 385-8777 if you do not receive all pages indicated above or experience any difficulty in receiving this facsimile.** | *This Facsimile is intended only for the use of the addressee and, if the addressee is a client or their agent, contains privileged and confidential information. If you are not the intended recipient of this facsimile, you have received this facsimile inadvertently and in error. Any review, dissemination, distribution, or copying is strictly prohibited. If you received this facsimile in error, please notify us by telephone and return the facsimile to us immediately.* |

PLEASE CONFIRM RECEIPT OF THIS TRANSMISSION BY FAXING A CONFIRMATION TO 972-385-7766.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: **Janakiraman et al.**   §   Group Art Unit: **2176**

§

Serial No.: **09/838,428**   §   Examiner: **Ries, Laurie Anne**

§

Filed: **April 19, 2001**   §   Attorney Docket No.: **AUS920010015US1**

§

For: **Displaying Text of Video in**
Browsers on a Frame by Frame Basis
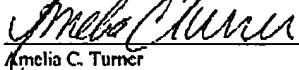
35525

PATENT TRADEMARK OFFICE
CUSTOMER NUMBER

TRANSMITTAL DOCUMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:
TRANSMITTED HEREWITH:

- Appeal Brief (37 C.F.R. 41.37).

A fee of $500.00 is required for filing an Appeal Brief. Please charge this fee to IBM Corporation Deposit Account No. 09-0447. No additional fees are believed to be necessary. If, however, any additional fees are required, I authorize the Commissioner to charge these fees which may be required to IBM Corporation Deposit Account No. 09-0447. No extension of time is believed to be necessary. If, however, an extension of time is required, the extension is requested, and I authorize the Commissioner to charge any fees for this extension to IBM Corporation Deposit Account No. 09-0447.

Respectfully submitted,
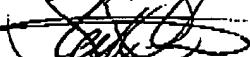
Stephen R. Tkacs
Registration No. 46,430
AGENT FOR APPLICANTS

Duke W. Yee
Registration No. 34,285
YEE & ASSOCIATES, P.C.
P.O. Box 802333
Dallas, Texas 75380
(972) 385-8777
ATTORNEY FOR APPLICANTS

Docket No. AUS920010015US1 *PATENT*

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: **Janakiraman et al.** §
§ Group Art Unit: **2176**
Serial No.: **09/838,428** §
§ Examiner: **Ries, Laurie Anne**
Filed: **April 19, 2001** §
§
For: **Displaying Text of Video in** §
Browsers on a Frame by Frame Basis §

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPEAL BRIEF (37 C.F.R. 41.37)

This brief is in furtherance of the Notice of Appeal, filed in this case on February 16, 2005.

The fees required under § 41.20(B)(2), and any required petition for extension of time for filing this brief and fees therefore, are dealt with in the accompanying TRANSMITTAL OF APPEAL BRIEF.

REAL PARTY IN INTEREST

The real party in interest in this appeal is the following party: International Business Machines Corporation.

RELATED APPEALS AND INTERFERENCES

With respect to other appeals or interferences that will directly affect, or be directly affected by, or have a bearing on the Board's decision in the pending appeal, there are no such appeals or interferences.

STATUS OF CLAIMS

A. TOTAL NUMBER OF CLAIMS IN APPLICATION

Claims in the application are: 1, 3-8, 10-15, and 17-21

B. STATUS OF ALL THE CLAIMS IN APPLICATION

1. Claims canceled: 2, 9, and 16
2. Claims withdrawn from consideration but not canceled: NONE
3. Claims pending: 1, 3-8, 10-15, and 17-21
4. Claims allowed: NONE
5. Claims rejected: 1, 3-8, 10-15, and 17-21
6. Claims objected to: NONE

C. CLAIMS ON APPEAL

The claims on appeal are: 1, 3-8, 10-15, and 17-21

STATUS OF AMENDMENTS

There are no amendments after final rejection.

SUMMARY OF CLAIMED SUBJECT MATTER

Independent claims 1, 8, and 15:

The presently claimed invention provides a method, computer program product, and system for presenting text from multimedia data to a user. The present invention receives multimedia data containing an associated plurality of sets of text data, wherein the plurality of sets of text data includes a first text data set associated with a first plurality of video frames and a second text data set associated with a second plurality of video frames. See specification, page 11, line 24, to page 12, line 15; page 16, line 26, to page 17, line 2. The present invention extracts the associated plurality of sets of text data from the multimedia data. See specification, page 11, lines 15-23; page 13, lines 23-31; page 14, line 25, to page 16, line 25; page 17, lines 3-12. The present invention outputs the first text data set with a one video frame of the first plurality and then, responsive to determining that the text in the multimedia data has changed from the first text data set to the second text data set, outputs the second text data set and a one video frame of the second plurality of video frames. See specification, page 11, line 24, to page 12, line 15; page 17, lines 12-22.

The means recited in independent claim 15, as well as dependent claims 17-21, may be data processing hardware within server 200, client 300, and combinations thereof, as described in the specification at page 6, line 2, to page 10, line 20, operating under control of software performing with the functionality described in the specification at page 10, line 21, to page 14, line 12, or equivalent.

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

The grounds of rejection on appeal are as follows:

Claims 1, 3-8, 10-15, and 17-21 are rejected under 35 U.S.C. § 103(a) as being allegedly unpatentable over *Gibbon et al.* (U.S. Patent Publication No. 2004/0078188 A1) in view of *Cramer et al.* (U.S. Patent Publication No. 2002/0104096 A1).

ARGUMENT

I. **35 U.S.C. § 103, Alleged Obviousness of claims 1, 3-8, 10-15, and 17-21**

The Final Office Action rejects claims 1, 3-8, 10-15, and 17-21 under 35 U.S.C. § 103(a) as being allegedly unpatentable over *Gibbon et al.* (U.S. Patent Publication No. 2004/0078188 A1) in view of *Cramer et al.* (U.S. Patent Publication No. 2002/0104096 A1). This rejection is respectfully traversed.

Gibbon teaches a system and method for automated multimedia content indexing and retrieval. *Gibbon* teaches separating a multimedia stream into audio, visual, and text components, segmenting the components based on semantic differences, identifying at least one target speaker, identifying a topic of the multimedia using the segmented text and topic category models, generating a summary of the multimedia event based on components, the identified topic, and the identified speaker, and generating a multimedia description of the multimedia event based on the target speaker, the identified topic, and the generated summary. See Abstract. *Gibbon* teaches that video and text are synchronized in time. A cited portion of *Gibbon* states:

> [0030] FIG. 1 shows an example of the content hierarchy of broadcast news for recovery. In this hierarchy, the lowest level contains the continuous multimedia data stream (audio, video, text). With the audio, video and text separated as shown **102**, linear information retrieval is possible. *The audio, video and text are synchronized in time.* Text may be from closed caption provided by a media provider or generated by the automatic speech recognition engine. If text originates from closed captioning, *time alignment between the audio and text needs to be performed.* At the next level, commercials are separated **104**. The remaining portion is the newscast **106**. The news is then segmented into the anchorperson's speech **108** and the speech from others **110**. The intention of this step is to use detected anchor's identity to hypothesize a set of story boundaries that consequently partition the continuous text into adjacent blocks of text. Higher levels of semantic units can then be extracted by grouping the text blocks into individualized news stories **112** and news introductions or summaries **114**. In turn, each news story can consist of either the story by itself or augmented by the anchorperson's introduction to the story. Using the extracted stories and summaries/introductions, topics can be detected and categorized **116**. The news content is thus finished as multimedia story content available for content-based browsing and nonlinear information retrieval **118**. Detailed

semantic structure at the story level is shown in **FIG. 2.** [emphasis added]

Gibbon, paragraph [0030]. *Gibbon* is primarily concerned with segmenting a multimedia stream so that it may be categorized and indexed.

In contradistinction, the present invention is concerned with presenting text to a user such that a first set of text is output with *one* video frame from a fist plurality associated with the first set of text data, and when text is changed to a second set of text, the second set of text is then output with *one* video frame from a second plurality associated with the second set of text data. The present invention recognizes the problem associated with text being synchronized in time with streaming media. The speed of the media stream, as well as the distractions presented by audio and video, may cause difficulties for people with visual and cognitive disabilities. The present invention solves this problem by presenting a set of text data with one video frame and then outputting a second set of text data, again presented with one video frame.

As acknowledged in the Final Office Action, *Gibbon* does not teach responsive to determining that the text in the multimedia data has changed from a first text data set to a second text data set, outputting the second text data set and a one video frame of the second number of video frames. The Final Office Action alleges that this feature is taught by *Cramer* in claim 15, which reads as follows:

> 15. A method of providing a web-based multimedia presentation on a remote user computer, comprising:
>> transmitting streamed content to the user computer for display within a first display screen of the web page; and
>> · transmitting video content to the user computer for display within a second display screen of the web page, wherein the video content includes embedded commands which control the display of the non-video content within the first display screen in synchronization with playing of the video content within the second display screen.

The method of *Cramer*, as detailed in the above claim, presents the problem of synchronizing display of a text-based document, a web page, with streamed content. Also, the method of *Cramer* presents streamed content, rather than one frame from a plurality of video frames. In contrast, the present invention solves the disadvantages of *Gibbon* and *Cramer* by outputting text data with a single frame of video data, rather than synchronizing text with a time-based moving video stream.

(Appeal Brief Page 9 of 19)
Janakiraman et al. – 09/838,428

PAGE 11/21 * RCVD AT 4/15/2005 4:12:49 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-1/7 * DNIS:8729306 * CSID:9723857766 * DURATION (mm-ss):05-24

In fact, *Cramer* teaches away from the claimed invention, because *Cramer* teaches transmitting streamed video content with commands to control the display of non-video content. That is, *Cramer* would lead a person of ordinary skill in the art to output multiple video frames with a set of text, rather than **one** video frame, as in the claimed invention. More particularly, *Cramer* would not lead a person of ordinary skill in the art to modify *Gibbon* to output a **one** video frame responsive to determining that the text in the multimedia data has changed from the first text data set to the second text data set.

Thus, *Gibbon* and *Cramer*, taken individually or in combination, fail to teach or suggest the claimed invention. Even if one were motivated to combine *Gibbon* and *Cramer* as proposed in the Final Office Action, the combination would not result in the claimed invention. Instead, a combination of *Gibbon* and *Cramer* would result in a method and system for indexing and summarizing multimedia content so that audio and video can be output in time synchronized fashion with text.

Moreover, the Examiner may not use the claimed invention as an "instruction manual" or "template" to piece together the teachings of the prior art so that the invention is rendered obvious. *In re Fritch*, 972 F.2d 1260, 23 U.S.P.Q.2d 1780 (Fed. Cir. 1992). Such reliance is an impermissible use of hindsight with the benefit of Appellants' disclosure. *Id.* Therefore, absent some teaching, suggestion, or incentive in the prior art, *Gibbon* and *Cramer* cannot be properly combined to form the claimed invention. As a result, absent any teaching, suggestion, or incentive from the prior art to make the proposed combination, the presently claimed invention can be reached only through an impermissible use of hindsight with the benefit of Appellants' disclosure a model for the needed changes. In fact, the Final Office Action actually states:

> Therefore, it would have been obvious to combine Cramer with Gibbon for the
> benefit of enabling the video to control the display video **to obtain the invention**
> as specified in claims 1, 8, and 15. [emphasis added]

Clearly, the rejection uses the claims as a blueprint to reconstruct the present invention.

Gibbon and *Cramer*, taken individually or in combination, do not teach or suggest each and every claim limitation. Therefore, *Gibbon* and *Cramer* do not render at least independent claims 1, 8, and 15 obvious. Since claims 3-7, 10-14, and 17-21 depend from claims 1, 8, and 15, the same distinctions between *Gibbon* and *Cramer* and claims 1, 8, and 15 apply for these claims. Furthermore, claims 3-7, 10-14, and 17-21 recite additional combinations of features not

(Appeal Brief Page 10 of 19)
Janakiraman et al. – 09/838,428

PAGE 12/21 * RCVD AT 4/15/2005 4:12:49 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-1/7 * DNIS:8729306 * CSID:9723857766 * DURATION (mm-ss):05-24

taught or suggested by the applied references.

Therefore, Appellants respectfully request that the rejection of claims 1, 3-8, 10-15, and 17-21 under 35 U.S.C. § 103(a) not be sustained.

I.A.   35 U.S.C. § 103, Alleged Obviousness of claims 3, 4, 10, 11, 17, and 18

With respect to claims 3, 10, and 17, the Final Office Action alleges that *Gibbon* teaches presenting more than one set of text data to the user simultaneously at paragraph [0030], which is reproduced above. Appellants respectfully disagree. As described in the cited portion and illustrated in the figures of the reference, *Gibbon* teaches that text from a multimedia stream is segmented into news stories. However, there is no teaching in *Gibbon* of receiving multimedia data containing a first set of text data associated with a first plurality of frames of video and a second set of text data associated with a second plurality of frames of video, outputting the first set of text data **with one video frame** of the first plurality of frames of video, and outputting the second set of text data **with one video frame** of the second plurality of frames of video **simultaneously with the first set of text data and the one video frame** from the first plurality. Rather, *Gibbon* teaches that users may browse the segmented news stories using a graphical user interface or table of contents and then play one news story or another with **streaming** video and audio. *Gibbon* does not teach or suggest presenting a first text data set **simultaneously** with a second text data set, wherein each text data set is output with a **one** video frame, as recited in the instant claims.

The applied references do not teach or suggest each and every claim limitation; therefore, *Gibbon* and *Cramer* do not render claims 3, 10, and 17 obvious. Since claims 4, 11, and 18 depend from claims 3, 10, and 17, the same distinctions between *Gibbon* and *Cramer* and claims 3, 10, and 17 apply for these claims.

I.A.1.   35 U.S.C. § 103, Alleged Obviousness of claims 4, 11, and 18

Furthermore, with respect to claims 4, 11, and 18, the Final Office Action alleges that *Gibbon* teaches presenting the first text data set and the second text data set simultaneously in separate frames in Figure 16 and at paragraph [0123]. The cited portions are as follows:

FIG. 16
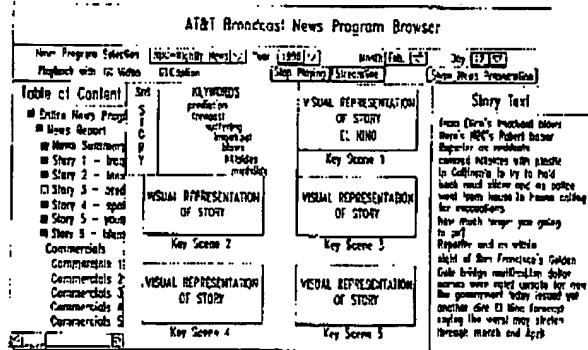
[0123] The exemplary representation of two stories are shown in FIGS. 16 and 17. The chosen stories are the third and fifth news program, respectively (which can be seen in the table of contents on the left portion of the interface). The representation for each story has three parts: the upper left corner is a set of 10 keywords automatically chosen from the segmented story text based on the relative importance of the words; the right part displays the full text of the story; the rest is the visual presentation of the story consisting of five images chosen from video in the content based manner described above.

The cited portions of *Gibbon* do not teach or suggest presenting a first text data set and a second text data set **simultaneously** in frames. Clearly, *Gibbon* teaches that either a first text data set is chosen and presented or a second text data set is chosen and presented, but not both simultaneously. Thus, the applied references do not teach or suggest each and every claim limitation, and *Gibbon* and *Cramer* do not render claims 4, 11, and 18 obvious.

I.B.   35 U.S.C. § 103, Alleged Obviousness of claims 7, 14, and 21

With respect to claims 7, 14, and 21, the Final Office Action alleges that *Gibbon* teaches extracting the number of sets of text data by parsing the multimedia data to determine the first text data set and the one video frame of the first number of video frames and discarding any moving image data in Figure 5, element 5040, and at paragraph [0037]. Element 5040 of Figure 5 is a flowchart step that states, "SEGMENT VIDEO, AUDIO, AND TEXT." Paragraph [0037] is as follows:

> [0037] In step 5040, the feature extraction unit 340 and the segmentation unit 350 identify features and parse the broadcast into segments. For example, separate news and commercials are

(Appeal Brief Page 12 of 19)
Janakiraman et al. – 09/838,428

PAGE 14/21 * RCVD AT 4/15/2005 4:12:49 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-1/7 * DNIS:8729306 * CSID:9723857766 * DURATION (mm-ss):05-24

> identified and segmented based on acoustic characteristics of audio
> data. FIGS. 6 and 7 show the typical waveforms for news reporting
> (FIG. 6) and commercials (FIG. 7). There is obviously a visual
> difference between the two waveforms. Such a difference is
> largely caused by the background music in the commercials. Thus,
> a set of audio features is adopted to capture this observed
> difference.

Nowhere does the cited portion, or any other portion, of *Gibbon* teach or suggest **discarding** moving image data. Rather, the cited portion of *Gibbon* describes feature extraction and segmentation, as well as waveforms for news reporting and commercials. The Final Office Action proffers no analysis as to how segmentation of multimedia into **video**, audio, and text or various audio waveforms are somehow equivalent to **discarding** moving image data, as recited in the instant claims. The applied references do not teach or suggest each and every claim limitation; therefore, *Gibbon* and *Cramer* do not render claims 7, 14, and 21 obvious.

II.   Conclusion

In view of the above, Appellants respectfully submit that claims 1, 3-8, 10-15, and 17-21 are allowable over the cited prior art and that the application is in condition for allowance. Accordingly, Appellants respectfully request the Board of Patent Appeals and Interferences to not sustain the rejections set forth in the Final Office Action.

Respectfully submitted,

Stephen R. Tkacs
Registration No. 46,430
YEE & ASSOCIATES, P.C.
PO Box 802333
Dallas, TX 75380
(972) 385-8777
Agent for Appellants

(Appeal Brief Page 13 of 19)
Janakiraman et al. – 09/838,428

PAGE 15/21 * RCVD AT 4/15/2005 4:12:49 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-1/7 * DNIS:8729306 * CSID:9723857766 * DURATION (mm-ss):05-24

CLAIMS APPENDIX

The text of the claims involved in the appeal reads:

1. A method for presenting text from multimedia data to a user, the method comprising:

receiving multimedia data containing an associated plurality of sets of text data, wherein the plurality of sets of text data includes a first text data set associated with a first plurality of video frames of the multimedia data, and a second text data set associated with a second plurality of video frames of the multimedia data;

extracting the associated plurality of sets of text data from the multimedia data;

outputting the first text data set with a one video frame of the first plurality of video frames; and

responsive to determining that the text in the multimedia data has changed from the first text data set to the second text data set, outputting the second text data set and a one video frame of the second plurality of video frames.

3. The method as recited in claim 1, wherein more than one of the plurality of sets of text data are presented to the user simultaneously.

4. The method as recited in claim 3, wherein the more than one of the plurality of sets of text data are presented in separate frames.

5. The method as recited in claim 1, wherein the first text data set and the second text data set are presented to the user individually in a sequential order.

(Appeal Brief Page 14 of 19)
Janakiraman et al. – 09/838,428

PAGE 16/21 * RCVD AT 4/15/2005 4:12:49 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-1/7 * DNIS:8729306 * CSID:9723857766 * DURATION (mm-ss):05-24

6.   The method as recited in claim 5, wherein a next set of text data in the sequential order is presented in response to an indication by the user to display the next set of text data.

7.   The method as recited in claim 1, wherein the step of extracting the plurality of sets of text data comprises parsing the multimedia data to determine the first text data set and the one video frame of the first plurality of video frames and discarding any moving image data.

8.   A computer program product in a computer readable media for use in a data processing system for presenting text from multimedia data to a user; the computer program product comprising:

first instructions for receiving multimedia data containing an associated plurality of sets of text data, wherein the plurality of sets of text data includes a first text data set associated with a first plurality of video frames of the multimedia data, and a second text data set associated with a second plurality of video frames of the multimedia data;

second instructions for extracting the associated plurality of sets of text data from the multimedia data;

third instructions for outputting the first text data set with a one video frame of the first plurality of video frames; and

fourth instructions that, responsive to determining that the text in the multimedia data has changed from the first text data set to the second text data set, output the second text data set and a one video frame of the second plurality of video frames.

(Appeal Brief Page 15 of 19)
Janakiraman et al. – 09/838,428

PAGE 17/21 * RCVD AT 4/15/2005 4:12:49 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-1/7 * DNIS:8729306 * CSID:9723857766 * DURATION (mm-ss):05-24

10.   The computer program product as recited in claim 8, wherein more than one of the plurality of sets of text data are presented to the user simultaneously.

11.   The computer program product as recited in claim 10, wherein the more than one of the plurality of sets of text data are presented in separate frames.

12.   The computer program product as recited in claim 8, wherein the first text data set and the second text data set are presented to the user individually in a sequential order.

13.   The computer program product as recited in claim 12, wherein a next set of text data in the sequential order is presented in response to an indication by the user to display the next set of text data.

14.   The computer program product as recited in claim 8, wherein the second instructions comprise instructions for parsing the multimedia data to determine the first text data set and the one video frame of the first plurality of video frames and discarding any moving image data.

15.   A system for presenting text from multimedia data to a user, the system comprising:

a receiver which receives multimedia data containing an associated plurality of sets of text data, wherein the plurality of sets of text data includes a first text data set associated with a first plurality of video frames of the multimedia data, and a second text data set associated with a second plurality of video frames of the multimedia data;

a text extraction unit which extracts the associated plurality of sets of text data from the multimedia data; and

an output unit which outputs the first text data set with a one video frame of the first plurality of video frames and, responsive to determining that the text in the multimedia data has changed from the first text data set to the second text data set, outputs the second text data set and a one video frame of the second plurality of video frames.

17.   The system as recited in claim 15, wherein more than one of the plurality of sets of text data are presented to the user simultaneously.

18.   The system as recited in claim 17, wherein the more than one of the plurality of sets of text data are presented in separate frames.

19.   The system as recited in claim 15, wherein the first text data set and the second text data set are presented to the user individually in a sequential order.

20.   The system as recited in claim 19, wherein a next set of text data in the sequential order is presented in response to an indication by the user to display the next set of text data.

21.   The system as recited in claim 15, wherein the extraction unit parses the multimedia data to determine the first text data set and the one video frame of the first plurality of video frames and discards any moving image data.

(Appeal Brief Page 17 of 19)
Janakiraman et al. – 09/838,428

PAGE 19/21 * RCVD AT 4/15/2005 4:12:49 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-1/7 * DNIS:8729306 * CSID:9723857766 * DURATION (mm-ss):05-24

EVIDENCE APPENDIX

There is no evidence to be presented.

(Appeal Brief Page 18 of 19)
Janakiraman et al. - 09/838.428

PAGE 20/21 * RCVD AT 4/15/2005 4:12:49 PM [Eastern Daylight Time] * SVR:USPTO-EFXRF-1/7 * DNIS:8729306 * CSID:9723857766 * DURATION (mm-ss):05-24

RELATED PROCEEDINGS APPENDIX

There are no related proceedings.